Dagmar Barth-Weingarten
(Hermann Paul School of Language Sciences, Freiburg, Germany)

The participant perspective: Interactional-linguistic work on the phonetics of talk-in-interaction

(Presentation at the panel "*Methods of analyzing spoken English*"
(convened by Neal Norrick)
at ISLE2, Boston, 17-21 June, 2011
Conference theme "Methods past and current")

(dagmar.barth-weingarten@hpsl.uni-freiburg.de)

Abstract
This paper presents the concept of the "participant perspective" as an approach to the study of spoken language. It discusses three aspects of this concept and shows that they can offer helpful tools in spoken language research. Employing the participant perspective provides us with an alternative to many of the approaches currently in use in the study of spoken language in that it favours small-scale, qualitative research that aims to uncover categories relevant for the participants. Its results can usefully complement large-scale studies of phenomena on all linguistic dimensions of talk.

1. Introduction[1]

The participant perspective (PP) is central to Interactional Linguistics (IL, see section 3). The latter is a relatively young research strand, which emerged from Conversation Analysis (CA, see, for instance, Hutchby/Wooffitt 1998, Schegloff 2007, Sidnell 2010) but focuses on linguistic phenomena. The concept of the PP has been adopted by IL from CA. I will first explain this concept, then summarize how it is employed in Interactional Linguistics and finally exemplify its application in the study of one linguistic phenomenon, namely prosodic-phonetic structuring.

2. The PP in CA

CA literature mentions the PP in connection with three different aspects:
1) The data are recorded, or video-taped, instances of everyday, natural, consequential talk.
This means that the talk studied should be talk that would have taken place anyway without the observer. Moreover, it should be talk that is consequential, i.e. what is being talked about, and how, has consequences, for the further course of talk and potentially also for the participants' lives. Thus, the language material studied should not have evolved in a lab situation. CA looks at the kind of language participants as members of a specific speech community encounter every day.
2) What is being looked for are emic categories (Pike 1967) and members' devices (Wootton 1989).
CA is interested in the organization and order of social action in interaction as relevant for the participants and accomplished by them (see Heritage 1984, Psathas 1995) because it understands "interaction as something sequentially and locally shaped between people in activity " (Martin 2004: 73). For this reason, it aims at reconstructing the "participant's own concepts and accounts" (Edwards/Stokoe 2004: 7), "the participants own handling and

[1] I would like to thank Margret Selting, Maxi Kupetz and the participants of the ISLE2 panel on "Methods of analyzing spoken English" for helpful comments on earlier versions of this presentation. Special thanks is due to Elizabeth Couper-Kuhlen for her insightful comments on the draft of this paper.

understanding of the system" (Seedhouse 2005), their "(ethno)means and methods" (Weber 2003).

3) Participants display (an orientation towards) orderliness in talk.

CA assumes that, in its contingencies, everyday talk is systematically organized and deeply ordered at all points (see, e.g., Heritage 1984). I.e. basically nothing happens in talk that is not systematic – I will come to so-called "speech errors" in a minute.

Moreover, this orderliness is used by the participants to accomplish actions and activities in interaction with each other. Therefore, the orderliness needs to be displayed by the participants to one other and it can therefore and in this way also be used by the analyst to detect patterns of language use (see Schegloff 1996a).

How does this work in practice? "The regularities are produced and oriented to by participants as normatively oriented-to grounds for inference and action" (Heritage 1984: 244). They are thus "made relevant in and through participants' action" (Martin 2004: 73, see also Pomerantz/Fehr 1997, Psathas 1995, Schegloff 1997, ten Have 1999). Hence, we need to study the participants' actions (Goodwin 1984: 243) to find out about the regularities of talk.

One phenomenon that is quite helpful in this respect is adjacency pair organization. The concept of adjacency pairs (AP) captures the observation that utterances usually come in pairs, such as

```
A:    Do you have a pen?
B:    Here you are.     vs.    Well, actually…
```

The production of a specific first pair part (FPP) makes the production of a specific second pair part (SPP) relevant next (see Schegloff 2007, also Levinson 1983). What this relevant SPP is, can be assessed on the basis of principles of sociality (see Schegloff 2007), experience (e.g., frequency of occurrence, lack of delay in turn-taking) etc. On the other hand, if it does not occur, the SPP is *noticeably* absent: it is either not produced at all or we don't find the expected SPP – it is either delayed (Pomerantz 1984) or the next move raises some problem or in some other way takes a course of action different from the expectable one etc. This means that the absence of an expectable SPP is observable and allows the participants to act accordingly.

This orderliness – the frequent occurrence of certain SPPs after certain FPPs, the production of unexpected SPPs and/or the problematization thereof – is also observable by the analyst. This can then be used as a basis for identifying 'normal' patterns of occurrence (of sequential organization, for instance) and how participants use them as well as for recognizing what is commonly referred to as "speech error", i.e. instances that for some reason deviate from the normal pattern. Against this background, APs can be used as a methodological tool: the next turn reveals whether we are dealing with an ordinary or a special instance of some pattern (see next turn proof procedure, Schegloff/Sacks 1973: 290, Hutchby/Wooffitt 1998, e.g.).

In fact, not only the 2nd but also the 3rd turn following some 1st can reveal orderliness: With smooth continuation, the original speaker ratifies the understanding of the 1st turn by the interlocutor as displayed in the 2nd turn. For an illustration, consider what we might want to call "garden-path parsing" of a FPP such as *Do you have a pen?* A simple *Yes*. in response to this would usually be problematized in the following SPP (e.g. *Well, I wanted you to give it to me.*) In this sense the next-turn proof procedure allows us to double-check whether we are dealing with a 'normal' instance of some sequence-organizational pattern; with some luck, the problematization may even tell us something about the pattern itself.

At the same time, this methodology requires us to look a single cases in their own right, at least at the beginning. Once a certain pattern has been described, more evidence can be gained from further instances, including so-called 'deviant' cases (Wootton 1989).

In sum, under the heading of the PP, CA practitioners investigate in a data-driven way

- individual instances of natural talk,
- in order to abstract from them the patterns (emic categories and devices) underlying talk
- by means of the orderliness of talk and the display thereof by the participants.

3. The PP in Interactional Linguistics

CA is a more sociologically oriented branch of research. Interactional Linguistics (IL) (see, for instance, Couper-Kuhlen/Selting 1996, 2001) has adopted its basic assumptions and methods and combined them with a more linguistic approach employing, for example, form-function pairings, syntagmatic and paradigmatic analysis and quantitative methods. In addition, IL incorporates concepts, methods and tools from compatible linguistic theories and models, e.g. Rhetorical Structure Theory, Empirical Construction Grammar and Genre Theory, Corpus Linguistics, Anthropological Linguistics among others.
With these, IL aims to study how participants employ linguistic resources to jointly accomplish (inter)action. This includes drawing inferences about what is going on and what is required from them at a particular point in interaction.
In doing so, IL employs the PP: Like CA, IL uses natural data and case-by-case analysis, it aims at discovering emic categories with regard to all aspects of language use on all levels of linguistic structure under the assumption of the orderliness of talk and its display in the data themselves.

A problem that becomes especially relevant in this approach to language is that linguistics is not an un-ploughed field. Linguists are professionals, trained in particular methods, in the framework of particular approaches, and these divide up perceptible reality in specific ways. Language is modeled in terms of certain categories and relevancies, most of which have not yet been checked against the PP.
Against this background, the issue of emic categories is especially challenging. While most of our categories and relevancies are motivated in and by *scholarly* frameworks and taxonomies, participants, first and foremost, aim to accomplish *action* in and through talk. These motivations do not necessarily always have the same effect. This is, of course, not to say that scholarly analysis is pointless altogether, but its assumptions and models are at least usefully checked against the PP. The range of findings on the systematicity of talk obtained by CA/IL work during the last 35 years (see, for instance, the references provided at http://www.paultenhave.nl/resource.htm) has shown how profitable the PP is. If we want to learn more about what enables participants to interact, we need to be able to take their perspective: the phenomena that are displayed as relevant and oriented to by them in everyday talk. As Schegloff (1996b) is eager to point out with regard to turn-constructional units, the components of turns:

> "[w]hat sorts of entities (described in grammatical or other terms) will be used and treated as turn-constructional units is determined by those who use the language […] not those who study it academically" (: 115)

In the remainder of my presentation I will show how the PP can be implemented as a tool for studying linguistic details of talk.

4. An example of implementing the PP in studying spoken language

In the framework of IL, the PP has been successfully employed with regard to syntactic theory (see, for instance, Auer's 2000, 2005 online syntax). As a sample phenomenon for this presentation, I have chosen prosodic-phonetic structuring of talk, because this is an issue

central to prosodic research. Moreover, there are theories on the interface of prosodic-phonetic structuring with syntax (e.g. Selkirk 2005) and information structure (e.g. Chafe 1994), so that this issue can be assumed to be of relevance for the study of (spoken) language in general.

4.1 Studying prosodic-phonetic structuring without the PP

Many earlier approaches to the prosodic-phonetic structuring of talk have tried to capture it in terms of the units perceivable. *Intonation group, tone group, intonation unit, declination unit, spoken paragraph* are just some of the labels in use (see, e.g., Palmer 1922, Crystal 1975, Brazil et al. 1980, Pierrehumbert 1980, Halliday 1985, Schuetze-Coburn et al. 1991, du Bois et al. 1992, Chafe 1994, Cruttenden 1997, Selting 2005). I will employ the label 'prosodic unit' (PU) for short. In the beginning, this model has presumably not been tested in terms of its interactional relevance for the participants, but it developed on the basis of the analysts' observation, and it has since then thriven in research on prosodic-phonetic structuring. Criteria that are recurrent in identifying PUs in the various approaches are:

- prominent syllables,
- coherent pitch contours,
- boundary tones and pitch steps,
- loudness and tempo changes,
- non-modal voice quality,
- pausing, and
- aspiration.

The concept of the PU has for the most part been used categorially. The existence of smaller and larger kinds of PU has been acknowledged, but they were still considered in terms of *units*, the 'bricks' of which an utterance is built. The autosegmental approach has drawn some attention to what comes *between* the units, the 'mortar', by assigning degrees of boundary strength (see Shattuck-Hufnagel/Turk 1996, for instance). Yet, their findings are mainly based on syntactic segmenting and at best read-aloud speech (the Boston news Corpus, see, for instance, Choi/Hasegawa-Johnson/Cole 2005). (There has been some empirical work on natural data by Nicole Dehé and Anne Wichmann in this framework, although they do not focus on the prosodic-phonetic structuring of talk in particular.)

The problem with a PU approach is that in natural talk – the first aspect of the PP – we find a considerable number of cases where it is difficult to identify PUs. Consider the following example by way of illustration. It is taken from the CallHome corpus, specifically from an AE telephone conversation between two nuns who are acquainted. Bonny had just recommended that Ann invite a travelling missionary priest to her convent and Ann is now in the process of suggesting what he could do there. The excerpt has been transcribed according to the GAT2 transcription conventions (see Selting et al. 2011). It is first displayed in the minimal transcript version and second as a fine transcript with prosodic-phonetic feature notated.[2]

(1) Sunday evening liturgy (CallHome, 4705_889-914) (minimal transcript)

```
1     Ann:  °hh and as you're speaking
2           i'm wondering about uhm (.) the idea of my certainly calling
            him
3           but also °hhh calling joanne sullivan
4           who is the campus minister at ro[semount.]
5     Bon:                                  [oh    ye]s
6           great
7     Ann:  °h and maybe working out
```

[2] For transcription conventions see Appendix. The relevant sound file will shortly be available at http://www.hpsl.uni-freiburg.de/barth-weingarten.

```
8           some kind of a deal whereby he comes and says (0.29) a sunday\
9           offers a sunday evening liturgy
10    Bon:     [oh:     ]      [(that's)      ]
11    Ann:  °hh[h (0.46)] and v[isits with the] students before or after
12    Bon:  (°h and see) he would love that
13          (0.3)
14    Bon:  he would love that
15          °hh let me just tell you
            ((continues with another anecdote illustrating that the
            missionary is well worth listening to))
```

(1') Sunday evening liturgy (CallHome, 4705_889-914) (fine transcript)

```
1     Ann:  °hh a:n' as ↑-YOU're ↓speaking-
2           =I'm wondering about uhm (.) |ʔ the idEa of:: |ʔ my ↑cErtainly
            calling -HIM-=
3           =but ↑Also °hhh calling joanne ↑-SUlli↓van-=
4           =who is the campus ↑MIni↓ste̲r̲ a̲t r̲o̲[se̲mo̲u̲nt.            ]
5     Bon:                                      [oh=<<cresc, dim>-YE:]S;>
6           =<<cresc, dim> -grEat;>
7     Ann:  °h an' mAybe working `OUT-
8           =↑sOme kind of a -dEal | wherebY he -cOm:es || an'
            -sAy:s || <<ʔ>(0.29)>_a -sUnday_((creak))\
9           ↑ʔoffers a -sUnda̲y̲ Evening
            ↑-LITur↓g<<l,dim>y̲:>;=
10    Bon:     [`Oh:;      ]       [(that's)      ]
11    Ann:  =°hh[h<<ʔ>(0.46)>]=↑an' v[Isits with the] -STUdents=|=before or
            <<dim>-AF↓te̲r̲:hh>.
12    Bon:  (°h an' see) ↑hE would -LOVE tha̲t.
13          (0.3)
14    Bon:  he would lOve ↓tha̲t.
15          =°hh ↑lEt me just -TELL you-
            ((continues with another anecdote illustrating that the
            missionary is well worth listening to))
```

The GAT2 notation attempts to capture the segmentation of speech in PUs: It requires the transcriber to put one PU ('intonation phrase' in GAT2 terminology) per line. We find many, indeed a majority of cases, where this can be done without any problem at all.
However, in lines 2 and 8 the segmentation is not so clear. In l. 2 we can notice some hesitation (*uhm*, micropause, lengthening). Perhaps due to that, the PU boundaries are less clearly marked. In line 8, however, there is no hesitation, and yet it is still difficult to determine where it should be segmented, especially when compared to the surrounding clearer marking, it could either be notated as one long PU until the repair at the end of l. 8, or – on the basis of the distribution of prominences, lengthening and rhythm – with PU endings after *comes*, *says* and perhaps even after *deal*.

To reconcile such cases with the PU approach, various solutions have been suggested in the past. They include:

- falling back on other linguistic dimensions, such as syntax (Cruttenden 1997, Crystal 1975), semantics, information structure (Chafe 1994), action (Szczepek Reed to appear)). Yet, depending on one's research focus, this may lead to circularity.
- proposing alternative prosodic criteria for segmentation, such as pauses (Brown et al. 1980). Yet these may be too loose since pauses may also occur within a PU (see, for instance, Couper-Kuhlen 1986).

- advocating a prototype approach to prosodic units (Brown et al. 1980, Gumperz/Berenz 1993, du Bois et al. 1992, du Bois 2008, Schuetze-Coburn 1992, 1994). Yet this is hard to handle in the actual practice of annotation.
- introducing lower-level prosodic units, such as the 'intonation subunit' (du Bois et al. 1992). Yet these have rarely been employed systematically.

Thus, all of these options involve potential difficulties, too. Moreover, regardless of whether one opts for dividing such an utterance into one or several IUs, a certain uneasiness remains (see Brown et al. 1980, Cruttenden 1997, du Bois et al. 1992 for a problematization of the PU), because no matter which way we go, the decision forces us to ignore a lot of potentially relevant information (see also Barnwell 2011). Most important of all: it feels untrue to what we perceive. Hence, none of these solutions is truly satisfactory.
In sum, the first aspect of the PP (natural data) has highlighted a problem with the established model of prosodic-phonetic structuring.
I will now show that other aspects of the PP can help solve the problem.

4.2 Prosodic-phonetic structuring in the PP

For this we need to go back to the basics. The first question when taking the PP as one's guide is not how to improve the analyst's model, but whether the concept we are dealing with is one that is relevant for the participants. So, what might participants need prosodic-phonetic structuring for? Auer (2010) suggests that it is for turn-taking: Interaction is an exchange of turns. Current speakers need to display possible turn endings to their interlocutors in order to accomplish interaction as smoothly as they in fact regularly do (see Sacks et al. 1974). If this is so, prosodic-phonetic structuring, or chunking, is very relevant and there are good chances of finding an orderly members' device here.
However, Auer also suggests that it is not the chunks themselves, but their endings that are relevant for the participants, because turn-taking happens at the end of utterances/units. So, what the participants need to look for is unit *endings*, their boundaries, not the units themselves. Consequently, Auer suggests shifting the focus of research to unit boundaries, or cesuras as I prefer to call them (see Barth-Weingarten 2011).
Hence, the second aspect of the PP (looking for emic categories) leads to an alternative approach to prosodic-phonetic structuring.

In addition, the PP helps us to investigate the prosodic-phonetic structuring of talk: If we look at cesuras from a PP, the second question is how to identify them in an objective and non-circular way without being influenced by the unit approach. This can be accomplished by employing the third aspect of the PP – the display of orderliness in talk and the participants' orientation towards it: We can look at cesuras that are identified as such by the participants themselves. The most visible evidence for orientation to prosodic-phonetic structuring is co-participants' behavior, specifically their 'incomings': According to Ford/Fox/Thompson (1996) "there is reason to believe that those syntactic boundaries *where speakers do come inn* [sic] will in fact be *prosodic boundaries* as well" (:429, my emphasis - DBW). Incomings include all sorts of verbal and non-verbal utterances, ranging from laughter, continuers, acknowledgements, assessments, requests for clarification, collaborative completions to full turns.[3] In the excerpt above we find instances of such incomings in lines 5-6, 10, 12 and 14. Once we have located these places, we can analyze them in order to study what methods the participants use to signal cesuras and consequently by what means the co-participants can recognize them.
For the actual prosodic-phonetic analysis we can fall back on the work of the York approach of Phonology/Phonetics for Conversation (see, e.g., Local/Kelly/Wells 1986, Kelly/Local

[3] In face-to-face conversation this also includes bodily behavior such as gazing and head nods.

1989, Local/Walker 2005). It, too, reconstructs linguistic features as members' devices in the organization and management of talk with a particular focus on prosodic-phonetic features (see also Couper-Kuhlen/Ford 2004, Barth-Weingarten/Reber/Selting 2010, for instance). The York school uses detailed parametric, impressionistic phonetic analysis as a method. In this way we can obtain a list of prosodic-phonetic cues that are potentially relevant for organizing incomings, because they occur at places in talk where the participants themselves orient towards speaker change.

In our example we find the following cues before the co-participant incomings:
In l. 5:

- a pitch peak (249Hz), followed by
- a long creaky stretch, which is perceptible as
- a pitch jump down across 29ST to the bottom of the speaker's voice range (47Hz) with the last syllable before the incoming being produced at 61Hz with a very short diplophonic, i.e. creaky, passage of 61/200Hz.

Before l. 10 we find:

- a pitch peak (203Hz), followed by
- a pitch jump down on the last syllable to below the middle of the speaker's voice range (156Hz), which towards the end of the syllable turns into
- a short creaky stretch almost at the bottom of the speaker's range (67Hz) (altogether 19ST),

and in addition there is

- loudness diminuendo and
- final lengthening, which is followed by
- a latched inbreath.

And before l. 12 we hear

- a pitch peak (180Hz), followed by
- a pitch fall on the beginning of the last syllable to below the middle of the speaker's voice range (166Hz), ending in
- a shorter creaky stretch at the very bottom of the speaker's range (55Hz) (altogether 20ST),

and in addition there is

- loudness diminuendo,
- final lengthening and
- an audible, lengthy aspirated release of the final segment.

First, we can note the overall parallelism in prosodic-phonetic marking before the incomings. This suggests that in this excerpt there is a recurrent set of specific features, which are used by the participant in cesura marking and which can serve the co-participant to recognize when incomings are relevant (members' devices). It includes a pitch peak followed by a noticeable downward pitch movement which ends at the bottom of the speaker's range and is perceptible as some kind of creak. Additional recurrent features are decreasing loudness and final lengthening. This set was identified by observing the participants' display of orderliness, the 3rd aspect of the PP.
Moreover it can be noted that the members of this list are, to a large extent, prosodic-phonetic features that have also elsewhere been considered as disintegrative (see, for instance, Local et al. 1986, Schönherr 1997, Birkner 2008). The list, in fact, parallels that of IU *boundary* features of the British School. To this we can, however, add two observations: First of all, the occurrence of a latched inbreath before the incoming in l. 10. It is interesting to find this feature, which has elsewhere been considered clearly projective (see Jefferson 1983a, for

instance), among the set of prosodic-phonetic features co-occurring with cesuras. It typically occurs at places of minimal response in my data (see also below). But further research is needed to show whether this is a cesuring feature typical of spoken language as produced in interaction – caused by on-line contingencies of talk, such as holding the floor. Second, and perhaps more importantly, we can add the observation that there are certain differences in the strength of boundary marking before the three incomings: Perceptually (and acoustically), the cesura before l. 12 is the strongest. It has the most extensive, i.e. largest, set of disintegrative features, the pitch falls lowest and the long final aspirated release clearly yields the floor. The set before l. 5 is perceptually almost as strong, although it neither exhibits that aspirated final release nor the final lengthening and the loudness diminuendo, or at least they are not audible to the co-participant before she starts coming in. Nevertheless, the speaker reaches the second lowest pitch level and for this covers a pitch range that is larger than that of the other incomings. In addition, the co-participants hears the longest stretch of creak. These features seem to be sufficient to indicate the relevance of an incoming, that can then even be produced with a recognitional onset (see Jefferson 1986). In comparison to these places of incoming, that before l. 10 is relatively weakly disintegrative. The pitch movement covers a wider range but does not fall to the lowest possible level, the creaky stretch is short, there is no audible release of the final segment, but instead a projective latched inbreath.

If we relate this difference in cesural strength to the kinds of interactional moves the participants are accomplishing with their incomings, we can moreover see an interesting correlation between the 'kind of response given' and the 'strength of cesura marking': The most extensive set of markings is followed by a full turn and a following topic shift in l. 12, 14-15 by the next speaker. Note also that the prior speaker (A) not only produces the extensive marking, but also passes the opportunity to continue topical talk in l. 13. Hence, both participants seem to orient to a turn-ending. The slightly less extensive set of cesural features before l. 5 is followed by less than a full turn - a short assessment – with which the prior speaker does not compete and after which she continues topical talk. Thus, here, too, both participants orient to the opportunity for an incoming, but with both participants there is evidence that that incoming is regularly less than a full turn. Finally, after the least disintegrative set of features the co-participant produces a subdued and hesitant news receipt *oh*, which is, with some delay, followed by what sounds like the aborted beginning of an assessment (*that's*, l. 10). As a result, the incoming is even less than a short assessment. In addition, this incoming occurs in overlap with that specific way of continuation of topical talk by the prior speaker that appears to be typical of places of minimal response at least in my data: In l. 11 Ann first produces an audible, lengthy in-breath followed by a holding pause (Local/Kelly 1986). This, in contrast with an outright verbal topical continuation (see l. 2, for instance), leaves space for an incoming. Note that in particular the holding pause functions as a specifically timed interactional resource here, that prolongs the space for incoming exactly until the *oh* is finished. Thus, both A and B orient to the occurrence of an incoming. At the same time, the in-breath and the holding pause project continuation, and the latter is produced by Ann at the first possible completion point of the incoming (see Local 1996). Moreover, Bonny does not compete with this continuation but aborts her move. Hence, both participants orient to the relevance of a *minimal* incoming.

These observations suggest that prosodic-phonetic cesuring in talk is in fact highly orderly: The less extensive the prosodic-phonetic cesural marking, the less turn-like the incoming (see also Barth-Weingarten 2009). Thus, cesuring in talk works gradiently and in some iconic way and it is functional in the organization of speaker change in interaction.

This impression is strengthened when we consider the places of speaker change discussed above against the background of those places where participants produce even less than a minimal response, i.e. where no response occurs. This is the case at the end of l. 3, for instance, where we find a pitch peak and subsequent jump down, but only across 6ST, and

where there is no other disintegrating feature. Here the co-participant produces no response. While this may be due to a lack of familiarity with the name mentioned, it is also striking that most of the prosodic-phonetic features preceding incomings are lacking and that the speaker continues without break. That is, neither of the participants orients to a cesura. A similar case, where the role of the missing prosodic-phonetic marking can be shown even more clearly, can be found in l. 11. There the turn is possibly complete after *visits with the students* but the current speaker continues smoothly till the next point of possible completion and again no incoming occurs. Arguably, the latter is the effect of the lack of cesuring prosodic-phonetic features. Hence, the orderliness we observed at places of speaker change also manifests itself at points in talk where no speaker change occurs, and both the current speaker and the potential next speaker orient towards it.

On the basis of these observations we conclude that prosodic-phonetic structuring should not only be framed in terms of PUs, nor even in terms of clear PUs and some fuzzy cases. Rather, we need to take into account the degree of cesuring and we can assume that it is systematic and functional, i.e., used as a members' device.
To be sure, variation in boundary marking has been postulated in the autosegmental approach, too, but I have arrived at it

- without retreat to other linguistic dimensions, which may introduce circularity,
- independently of the analyst's interpretation of a unit and its ending,
- by examining natural talk and by doing justice to the single case, and
- by incorporating participants' needs and methods,

that is, by employing the PP.

Finally, if we think of the degree of cesural marking as a participant's device, we can also deal with fuzzy cesuras: Looking at prosody and phonetics in its own right, instead of having recourse to other linguistic dimensions, allows us to take note of the things that happen at places like fuzzy cesuras not as something problematic, but as something potentially interesting. Thus, to return to l. 8 in our excerpt, the question would not be whether it is one or several units, but which cesural markers are there, to what extent and why. If we notate these fuzzy boundaries as candidate prosodic cesuras, we can later return to those places and study in detail what happens there in terms of prosody and phonetics in order to try to find out what fuzzy cesuras might be used for by the participants (see Barth-Weingarten in prep.). This kind of method is not without challenge when it comes to its actual implication, yet the insights to be gained are worth the effort.

5. Summary and conclusions

This paper presented the participant perspective (PP) as it is employed in CA and IL and argued that it is a useful approach to studying spoken language. Three aspects of the PP have been discussed and their employment has been illustrated with the study of one linguistic aspect of talk, prosodic-phonetic structuring.
By using natural data and looking for emic categories and members' devices on the basis of their orderly use

- the problems of the earlier theoretical model were highlighted (i.e., the unit approach in natural talk),
- an alternative, more objective and less potentially circular approach was developed, one that takes into account the participants' needs and which solves a number of the problems of previous models (the cesura approach),
- a tool for studying cesuring in talk was provided ((co-)participants' behavior), and
- a new area of research opened up (fuzzy prosodic-phonetic cesuras).

A PP approach has thus proved to be a fruitful and feasible methodological alternative for investigating spoken-language phenomena in interaction. The findings that have been published by CA and IL during the last 35 years on the fine linguistic details of talk are evidence to its wide and successful applicability and put these approaches on a par with other methods for studying spoken language, such as those employed by large-scale corpus-based studies.

References

Auer, P. (2000): On line-Syntax – Oder: was es bedeuten könnte, die Zeitlichkeit der mündlichen Sprache ernst zu nehmen. In: Sprache und Literatur 85, 43-56.

Auer, P. (2005): Projection in interaction and projection in grammar. In: Text 25 (1), 7-36.

Auer, P. (2010): Zum Segmentierungsproblem in der gesprochenen Sprache. In: InLiSt - Interaction and Linguistic Structures, No. 49, November 2010, URL: http://www.inlist.uni-bayreuth.de/issues/49/InLiSt49.pdf (last access: 06.01.2011).

Barnwell, B. (2011): The perceptual basis of intonation units. (manuscript, Febr. 14, 2011).

Barth-Weingarten, D. (2009): "When to say something - some observations on prosodic-phonetic cues to the placement and types of responses in multi-unit turns." In: Barth-Weingarten, D./ N. Dehé/ A. Wichmann (eds.): Where prosody meets pragmatics. Bingley: Emerald, 143-182.

Barth-Weingarten, D. (2011): The fuzziness of intonation units: Some theoretical considerations and a practical solution. In: InLiSt – Interaction and Linguistic Structure, No. 51, May 2011, URL: http://www.inlist.uni-bayreuth.de/issues/51/index.htm (last access: 07.06.2011).

Barth-Weingarten, D. (in prep.): Cesuring in talk in interaction. Post-doc thesis, Freiburg University.

Barth-Weingarten, D./ E. Reber/ M. Selting (eds.) (2010): Prosody in interaction. Amsterdam: Benjamins.

Birkner, K. (2008): Relativ(satz)konstruktionen im gesprochenen Deutsch. Syntaktische, prosodische, semantische und pragmatische Aspekte. Berlin, New York: de Gruyter.

Brazil, D./ M. Coulthard/ C. Johns (1980): Discourse Intonation and Language Teaching. London: Longman.

Brown, G./ K.L. Currie/J. Kenworthy (1980): Questions of intonation. London: Croom Helm.

Chafe, W. (1994): Discourse, consciousness and time. The flow and displacement of conscious experience in speaking and writing. Chicago: Chicago University Press.

Choi, J-Y./ M. Hasegawa-Johnson/ J. Cole (2005): Finding intonational boundaries using acoustic cues related to the voice source. In: Journal of the Acoustical Society of America 118 (4), 2579-2587.

Couper-Kuhlen, E. (1986): An Introduction to English prosody. Tübingen: Niemeyer.

Couper-Kuhlen, E./ C. E. Ford (eds.) (2004): Sound patterns in interaction. Amsterdam: Benjamins.

Couper-Kuhlen, E./ M. Selting (1996): Towards an interactional perspective on prosody and a prosodic perspective on interaction. In: Couper-Kuhlen, E./ M. Selting (eds.): Prosody in Conversation: Interactional Studies. Cambridge: CUP, 11-56.

Couper-Kuhlen, E./ M. Selting (2001): Introducing interactional linguistics. In: Selting, M./ E. Couper-Kuhlen (eds.): Studies in interactional linguistics. Amsterdam: Benjamins, 1-22.

Cruttenden, A. (1997): Intonation. Cambridge: CUP.

Crystal, D. (1975): The English tone of voice. Essays in intonation, prosody and paralanguage. London: Arnold.

Du Bois, J.W. (2008): Rhythm and tunes: The intonation unit in the structure of dialogic engagement. Pres. held at the International Conference “Prosody and interaction”, Potsdam, 15-17-. September 2008.

Du Bois, J.W./ S. Schuetze-Coburn/ D. Paolino/ S. Cummings (1992): Discourse transcription, Santa Barbara Papers in Linguistics 4, Dept. of Linguistics, University of California, Santa Barbara.

Edwards, D./ E. Stokoe (2004): Commentary: Discursive psychology, focus group interviews and participants' categories. In: British Journal of Developmental Psychology 22, 499-507.

Ford, C.E./ B.A. Fox/ S.A. Thompson (1996): Practices in the construction of turns: The "TCU" revisited. In: Pragmatics 6, 427-454.

Goodwin, C. (1984): Notes on story structure and the organization of participation. In: Atkinson, J. M./ J. Heritage (eds.): Structures of Social Action. Studies in Conversation Analysis. Cambridge: CUP, 225-246.

Gumperz, J./ N. Berenz (1993): Transcribing conversational exchanges. In: Edwards, J./ M.D. Lampert (eds.): Talking data. Transcription and coding in discourse research. Hillsdale, N.Y.: Erlbaum, 91-121.

Halliday, M.A.K. (1985a): An introduction to functional grammar. London: Arnold.

Heritage, J. (1984): Garfinkel and ethnomethodology. Cambridge: Polity Press.

Hutchby, I./ R. Wooffitt (1998): Conversation analysis. Principles, practices and applications. Cambridge: Polity Press.

Jefferson, G. (1983a): Notes on some orderliness of overlap onset. In: D'Urso, V./ P. Leonardi (eds.): Discourse analysis and natural rhetorics. Padua: Cleup Editors, 11-38.

Jefferson, G. (1986): Notes on 'latency' in overlap onset. In: Human Studies 9, 153-183.

Kelly, J./ J. Local (1989): Doing phonology. Observing, Recording, Interpreting. Manchester, New York: Manchester University Press.

Levinson, S. (1983): Pragmatics. Cambridge: CUP.

Local, J. (1996): Conversational phonetics: some aspects of news receipts in everyday talk. In: Couper-Kuhlen, E./ M. Selting (eds.): Prosody in conversation. Cambridge: CUP, 177-230.

Local, J./ J. Kelly (1986): Projection and 'silences': Notes on phonetic and conversational structure. In: Human Studies 9, 185-204.

Local, J./ J. Kelly/ W.H.G. Wells (1986): Towards a phonology of conversation: turn-taking in Tyneside English. In: Journal of Linguistics 22, 411-437.

Local, J./ G. Walker (2005): Methodological imperatives for investigating the phonetic organization and phonological structures of spontaneous speech. In: Phonetica 62, 120-130.

Martin, C. (2004): From other to self. Learning as Interactional Change. Acta Universitatis Upsaliensis. Uppsala.

Palmer, H. (1922): English intonation, with systematic exercises. Cambridge : Heffer.

Pierrehumbert, J.B. (1980): The phonetics and phonology of English intonation. PhD, MIT.

Pike, K. L. (1967): Language in relation to a unified theory of the structure of human behaviour. 2nd edition. The Hague: Mouton.

Pomerantz, A. (1984): Agreeing and disagreeing with assessments: some features of preferred/dispreferred turn shapes. In: Atkinson, J.M./ J. Heritage (eds.): Structures of Social Action. Studies in Conversation Analysis. Cambridge: CUP, 57-101.

Pomerantz, A./ B.J. Fehr (1997) Conversation Analysis: An Approach to the Study of Social Action as Sense Making Practices. In: van Dijk, T. A. (ed.): Discourse as Social Interaction. London: Sage, 64-91.

Psathas, G. (1995): Conversation Analysis. Thousand Oaks, CA: Sage.

Sacks, H./ E.A. Schegloff/ G. Jefferson (1974): A simplest systematics for the organization of turn-taking for conversation. In: Language 50, 696-735.

Schegloff, E.A. (1996a): Confirming allusions: Toward an empirical account of action. In: American Journal of Sociology 102, 161-216.

Schegloff, E.A. (1996b): Turn-organization: one intersection of grammar and interaction. In: Ochs, E./ E. Schegloff/ S.A. Thompson (eds.): Interaction and grammar. Cambridge: CUP, 52-133.

Schegloff, E.A. (1997): Practices and actions: Boundary cases of other-initiated repair. In: Discourse Processes 23, 499-545.

Schegloff, E.A. (2007): Sequence Organization in interaction. A primer in conversation analysis (Vol. 1). Cambridge: CUP.

Schegloff, E. A./ H. Sacks (1973): Opening up closings. In: Semiotica 7: 289-327.

Schönherr, B. (1997): Syntax – Prosodie – nonverbale Kommunikation. Empirische Untersuchungen zur Interaktion sprachlicher und parasprachlicher Ausdrucksmittel im Gespräch. Tübingen: Niemeyer.

Schuetze-Coburn, S. (1992): Prosodic phrase as prototype. In: Proceedings of the IRCS workshop on prosody in natural speech. Vol. 1, Institute for Research in Cognitive Science, Univ. of Pennsylvania, Philadelphia, PA, 171-180.

Schuetze-Coburn, S. (1994): Prosody, syntax, and discourse pragmatics: Assessing information flow in German conversation. PhD dissertation, University of California, Los Angeles.

Schuetze-Coburn, S./ M. Shapley/ E.G. Weber (1991): Units of intonation in discourse: A comparison of acoustic and auditory analyses. In: Language and speech 34, 207-234.

Seedhouse, P. (2005): Conversation analysis and language learning. In: Language Teaching 38, 165-187.

Selkirk, E. (2005): Comments on intonational phrasing in English. In: Frota, S./ M. Vigario/ M.J. Freitas (eds.): Prosodies. With special reference to Iberian languages. Berlin: Mouton de Gruyter, 11-58.

Selting, M. (2005): Syntax and prosody as methods for the construction and identification of turn-constructional units in conversation. In: Hakulinen, A./ M. Selting (eds.): Syntax and Lexis in conversation. Studies on the use of linguistic resources in talk-in-interaction. Amsterdam, Philadelphia: Benjamins, 17-44.

Selting, M./ P. Auer/ D. Barth-Weingarten/ J. Bergmann/ P. Bergmann/ K. Birkner/ E. Couper-Kuhlen/ A. Deppermann/ P. Gilles/ S. Günthner/ M. Hartung/ F. Kern/ C. Mertzlufft/ C. Meyer/ M. Morek/ F. Oberzaucher/ J. Peters/ U. Quasthoff/ W. Schütte/ A. Stukenbrock/ S. Uhmann (in print, 2011): A system for transcribing talk-in-interaction: GAT 2 (translated and adapted for English by Elizabeth Couper-Kuhlen and Dagmar Barth-Weingarten). In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion (www.gespraechsforschung-ozs.de).

Shattuck-Hufnagel, S./ A.E. Turk (1996): A Prosody Tutorial for Investigators of Auditory Sentence Processing In: Journal of Psycholinguistic Research 25 (2), 193-247.

Sidnell, J. (2010): Conversation Analysis. An introduction. Chichester: Wiley and sons.

Szczepek Reed, B. (to appear): Prosody, syntax and action formation: Intonation phrases and 'action components'. In: Bergmann, P./ J. Brenning/ M. Pfeiffer/ E. Reber (eds.): Prosody and visual signals. Approaches towards usage-based grammar theories. (series linguae et litterae). Berlin, New York: de Gruyter.

ten Have, P. (1999): Doing conversation analysis. A practical guide. London: Sage.

Weber, T. (2003): The is no objective subjectivity in the study of social interaction. In: Forum Qualitative Sozialforschung / Forum: Qualitative Social Research [On-line Journal] 4, http://www.qualitative-research.net/fqs-texte/2-03/2-03weber-e.htm. [last access: 14.01.2006].

Wootton, A. J. (1989): Remarks on the methodology of conversation analysis. In. Roger, D./ P. Bull (eds.): Conversation. Clevedon: Multilingual matters, 238-258.

Appendix: Transcription conventions

GAT2 (Selting et al. 2011)

Minimal transcript

Sequential structure

`[ ]` `[ ]`	overlap and simultaneous talk

In- and outbreaths

`°h / h°`	in- / outbreaths of appr. 0.2-0.5 sec. duration
`°hh / hh°`	in- / outbreaths of appr. 0.5-0.8 sec. duration
`°hhh / hhh°`	in- / outbreaths of appr. 0.8-1.0 sec. duration

Pauses

`(.)`	micro pause, estimated, up to 0.2 sec. duration appr.
`(-)`	short estimated pause of appr. 0.2-0.5 sec. duration
`(--)`	intermediary estimated pause of appr. 0.5-0.8 sec. duration
`(---)`	longer estimated pause of appr. 0.8-1.0 sec. duration
`(0.5)/(2.0)`	measured pause of appr. 0.5/ 2.0 sec. duration (to tenth of a second)

Other segmental conventions

`and_uh`	cliticizations within units
`uh, uhm, etc.`	hesitation markers, so-called "filled pauses"

Laughter and crying

`haha, hehe, hihi`	syllabic laughter
`((laughs)), ((cries))`	description of laughter and crying
`<<laughing> >`	laughter particles accompanying speech with indication of scope
`<<:-)> so>`	smile voice

Continuers

`hm, yes, no, yeah`
`hm_hm, ye_es,`
`no_o`
	monosyllabic tokens
	bi-syllabic tokens

`ʔhmʔhm`	with glottal closure, often negating

Other conventions

`((coughs))` `<<coughing> >`	non-verbal vocal actions and events with indication of scope
`( )`	unintelligible passage
`(xxx), (xxx xxx)`	one or two unintelligible syllables
`(may i)`	assumed wording
`(may i say/let us say)`	possible alternatives
`((unintelligible, appr. 3 sec))`	unintelligible passage with indication of duration
`((...))`	omission in transcript
→	refers to a line of transcript relevant in the argument

Basic transcript

Sequential structure

=	fast, immediate continuation with a new turn or segment (latching)

Other segmental conventions

:	lengthening, by about 0.2-0.5 sec.
::	lengthening, by about 0.5-0.8 sec.
:::	lengthening, by about 0.8-1.0 sec.
?	cut-off by glottal closure

Accentuation

SYLlable	focus accent
!SYL!lable	extra strong accent

Final pitch movements of intonation phrases

?	rising to high
,	rising to mid
–	level
;	falling to mid
.	falling to low

Other conventions

<<surprised> >	interpretive comment with indication of scope

Fine Transcript

Accentuation

SYLlable	focus accent
sYllable	secondary accent
!SYL!lable	extra strong accent

Pitch jumps

↑	smaller pitch upstep
↓	smaller pitch downstep
↑↑	larger pitch upstep
↓↓	larger pitch downstep

Changes in pitch register

<<l> >	lower pitch register
<<h> >	higher pitch register

Intralinear notation of accent pitch movements

`SO	falling
´SO	rising
¯SO	level
^SO	rising-falling
ˇSO	falling-rising
↑`	small pitch upstep to the peak of the accented syllable
↓´	small pitch downstep to the valley of the accented syllable
↑¯SO bzw. ↓¯SO	pitch jumps to higher or lower level accented syllables
↑↑`SO bzw. ↓↓´SO	larger pitch upsteps or downsteps to the peak or valley of the accented syllable

Loudness und tempo changes, with scope

<<f> >	forte, loud
<<ff> >	fortissimo, very loud
<<p> >	piano, soft
<<pp> >	pianissimo, very soft
<<all> >	allegro, fast
<<len> >	lento, slow
<<cresc> >	crescendo, increasingly louder
<<dim> >	diminuendo, increasingly softer

`<<acc>  >`	accelerando, increasingly faster
`<<rall>  >`	rallentando, increasingly slower

Changes in voice quality and articulation, with scope

`<<creaky>  >`	glottalized
`<<whispery>  >`	change in voice quality as stated

Additional transcription symbols

|	fuzzy cesura (weak)
||	fuzzy cesura (stronger)
↑	small pitch upstep
↓	small pitch downstep
e̯	glottalization of the phoneme stated